The Voice Inside Your Head: Literally

Echoes of Thoughts !

Where Machines Go From "What Is This?" to "I Know What You're Thinking!"

Prepare yourself for an exhilarating challenge that will take you on a learning adventure, exposing you to the intricate and demanding nature of real-world pipelines. Be ready for a challenge, as the data type you'll be working with is renowned for its complexity. Approach with care and get set to enhance your skills !

Instructions (Read With Your Inner Voice)

- You're allowed to use **ANY** model, training paradigm, or neural architecture your brilliant mind can conjure up!
- Pick any training metric that tickles your fancy - but remember, final evaluations will only be based on **SSIM Index** and **Multi-Scale Spectral MSE Loss**.
- The use of pre-trained models for any sub-task or task is strictly prohibited. Violation of this rule will result in a score of **0 marks** for the entire competition.
- You will be competing against one another. There are ∞ possible methods for developing your solution's **architecture, optimizer, and training paradigm**. If three or more submissions are found to have architectures that are **60% or more similar**, those submissions will receive **0 marks** and may be reported to DAC for plagiarism. (Note: Similarity testing will be automated.)
- If the complete pipelines of any two submissions are found to be similar, both will receive **0 marks** and will be reported to DAC for plagiarism.
- You must submit your code using the provided stencil.

Dataset

- The provided EEG dataset includes high-resolution recordings of the brain's electrical activity, captured via a 64-channel EEG system following the 10-20 international system. The data is sampled at 1000 Hz and stored in two .pth files, containing full EEG recordings for each of 16 subjects.
- During the experiment, participants were shown images corresponding to distinct object classes for a duration of 0.5 seconds. Following the presentation of each image, the participant was asked to vocalize the class label of the object they had seen. This verbal task, combined with the EEG recording, allows for an analysis of both visual processing and cognitive response in the brain. The neural data captured during this activity provides valuable information on how the brain processes visual stimuli and generates motor responses (i.e., speech production).
- In summary, the dataset consists of synchronized EEG signals, capturing the brain's response to visual stimuli and cognitive tasks. The .pth files contain raw EEG data and labels useful for brain-computer interface applications and cognitive state analysis.

The Task: Create a Mind Listener

- Develop a deep learning model that processes EEG signals to extract what participants are thinking and generate spectrogram for the spoken audio.
- Each sample in your dataset is a dictionary with following keys:
  - `eeg_data`: EEG recording for ≈0.5 sec of shape (64,seq_len). 64 channels are ordered as:

```
eeg_channels = [Fp1, Fpz, Fp2, AF7, AF3, AFz, AF4, AF8,
                F7, F5, F3, F1, Fz, F2, F4, F6, F8,
                FT7, FC5, FC3, FC1, FCz, FC2, FC4, FC6,
                FT8, T7, C5, C3, C1, Cz, C2, C4, C6,
                T8, TP7, CP5, CP3, CP1, CPz, CP2, CP4,
                CP6, TP8, P7, P5, P3, P1, Pz, P2, P4,
                P6, P8, PO7, PO3, POz, PO4, PO8,
                O1, Oz, O2, FT9, FT10, TP9, TP10
                ]
```

  - `granularity`: EEG electrode placement
  - `subject`: Id of the participant
  - `label`: Label of the class
  - `image`: Image path of the class
- In a separate file `file_name` mapping of class_id to class name is given.

Evaluation Criteria & Submissions

- **Preparation Period:** You have 4 days to understand the dataset, develop a plan, and assess the problem's criticality.
- **Submission Phase:** After the preparation, you can submit one solution per day via a designated website (link provided at the start).
- **Results & Leaderboard:** Results will be emailed within a day of submission, with no public leaderboard; the final private leaderboard will be shared at the competition's end.
- **Submission:**
  - **Test Set Details:** The test set contains 12 EEG samples that are entirely distinct from the training data provided to you. These samples must be processed and evaluated independently.
  - **Task:** You need to generate a log-mel spectrogram from each of the 12 EEG signals. The resulting spectrograms must be saved in .pth format, as shown in the stencil code example.
  - **Submission Requirements:** After generating the log-mel spectrograms for all 12 samples and there respective audio, you must zip all 12 .pth files and corresponding audios and upload them to the competition website.
  - **Evaluation Metrics:** The competition uses SSIM (Structural Similarity Index) and Multi-Scale Spectral MSE Loss as evaluation metrics. Ensure that your code adheres to the specified constraints in the stencil code, as any deviation, even in minor details like precision, may result in negative outcomes.
  - **Evaluation Eligibility:** Your generated audio will be evaluated before computing performance metrics. A viva and code demonstration will be held at the end of the competition, and final marks will be based on both. For the demo, save all learnable parameters offline and bring the complete code and saved model.